Jekyll RDF

Template-Based Linked Data Publication with Minimized Effort and Maximum Scalability

Natanael Arndt <https://natanael.arndt.xyz/#i>

2019-06-13

ICWE 2019, Daejeon, South Korea

funded by

DAAD Deutscher Akademischer Austauschdienst
German Academic Exchange Service

Problem

Problem

RDFa 1.1 Primer: `https://www.w3.org/TR/xhtml-rdfa-primer/`

This is what humans see.

Problem

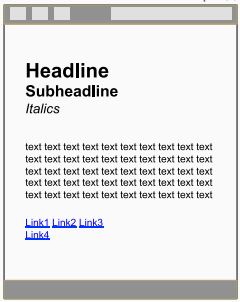
RDFa 1.1 Primer: `https://www.w3.org/TR/xhtml-rdfa-primer/`

This is what humans see.

Problem

RDFa 1.1 Primer: https://www.w3.org/TR/xhtml-rdfa-primer/

This is what browsers see.

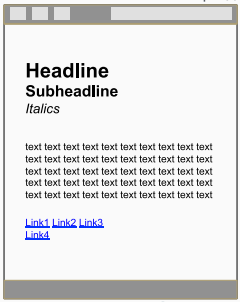


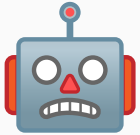
This is what humans see.

Problem

RDFa 1.1 Primer: https://www.w3.org/TR/xhtml-rdfa-primer/

This is what browsers see.



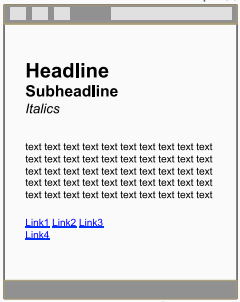


This is what humans see.

Problem

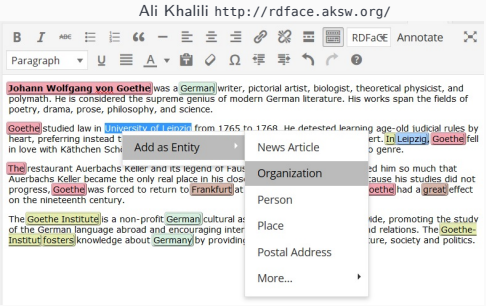
RDFa 1.1 Primer: https://www.w3.org/TR/xhtml-rdfa-primer/

This is what browsers see.


Headline
Subheadline
Italics
text text text

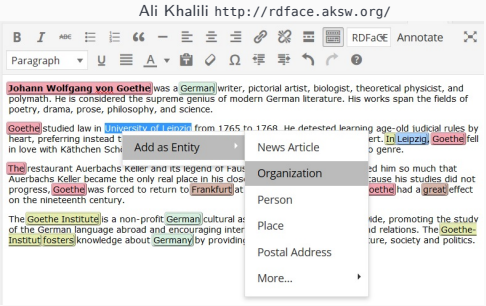
Problem

Yes, RDF and other formats help us to annotate and express data on the Web.



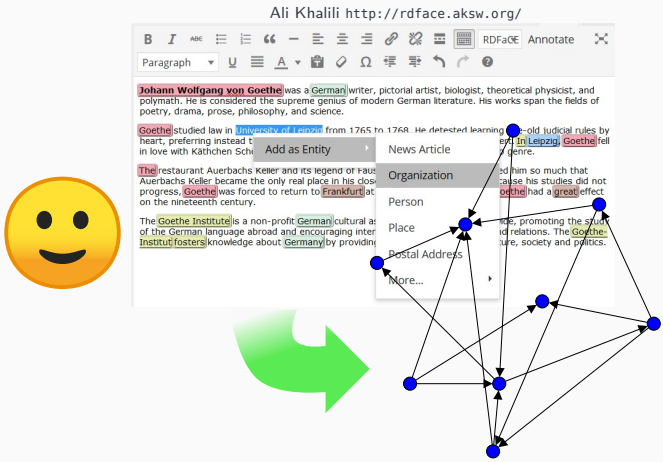
Problem

Yes, RDF and other formats help us to annotate and express data on the Web.



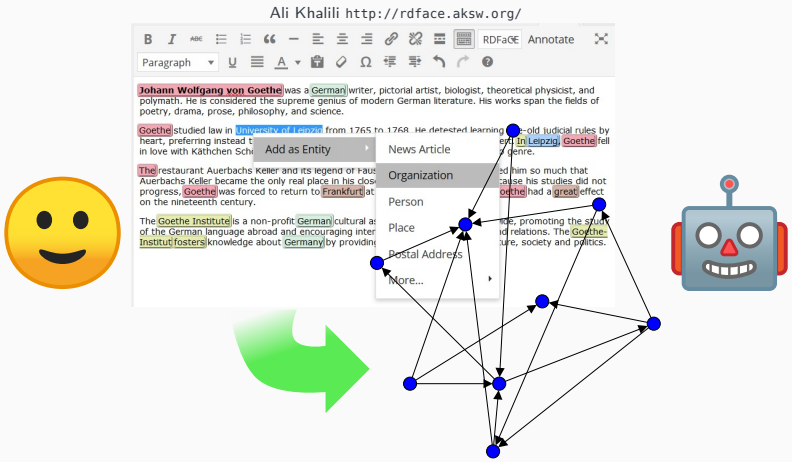
Problem

Yes, RDF and other formats help us to annotate and express data on the Web.



Problem

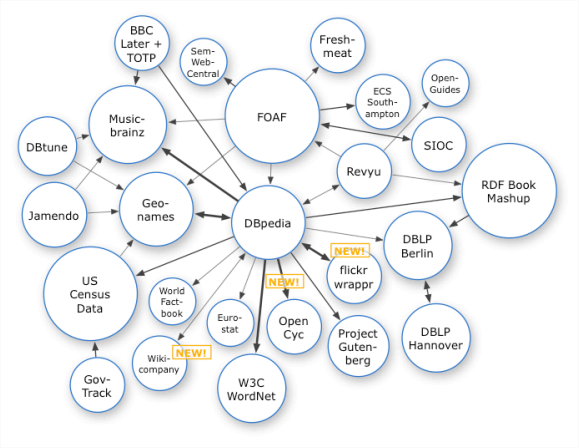
Yes, RDF and other formats help us to annotate and express data on the Web.

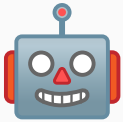


Problem

An increasing amount of data is published as RDF resp. Linked Data ...

The Linked Open Data Cloud https://www.lod-cloud.net/

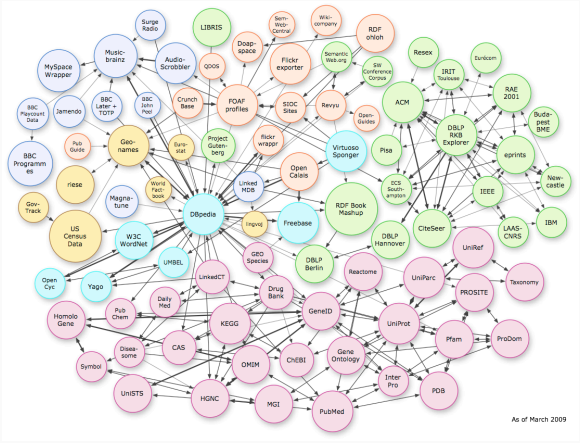


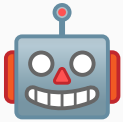


Problem

An increasing amount of data is published as RDF resp. Linked Data ...

The Linked Open Data Cloud `https://www.lod-cloud.net/`

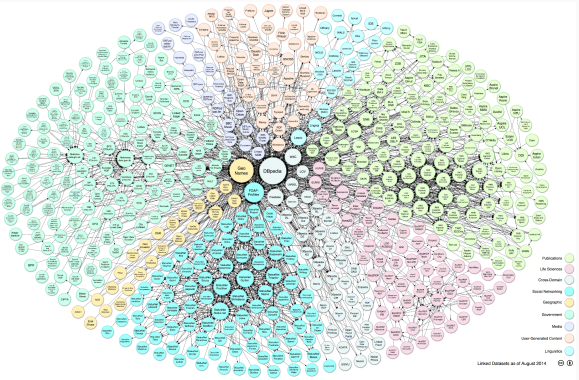




Problem

... while this data is hardly accessible to humans.

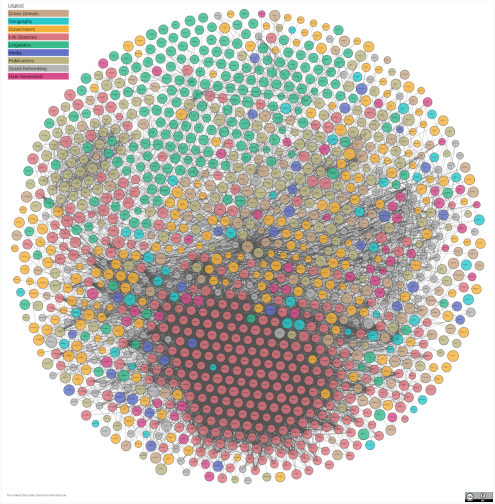
The Linked Open Data Cloud https://www.lod-cloud.net/

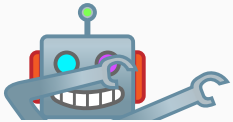


Problem

... while this data is hardly accessible to humans.

The Linked Open Data Cloud https://www.lod-cloud.net/

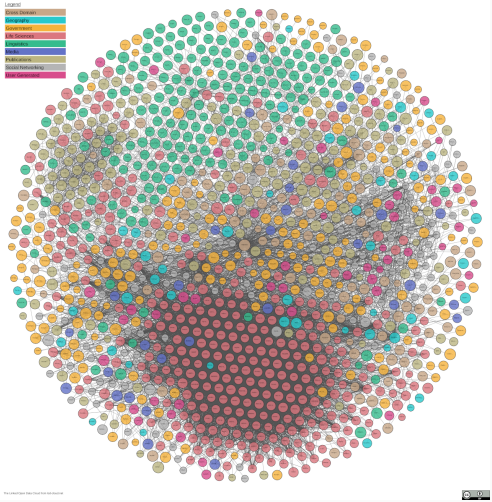


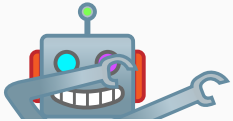


Problem

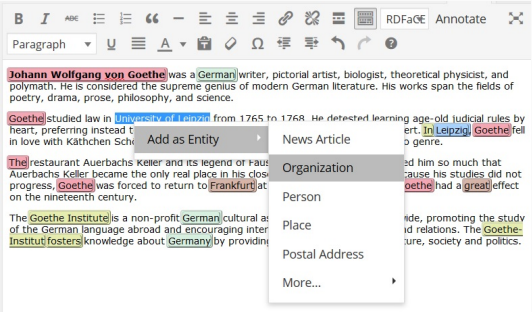
How can we make Data on the Web accessible to machines and to humans?

The Linked Open Data Cloud https://www.lod-cloud.net/

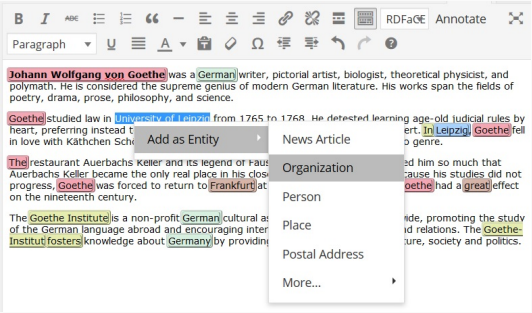


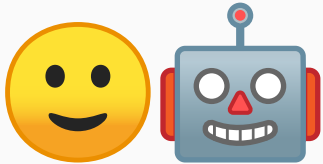


Problem

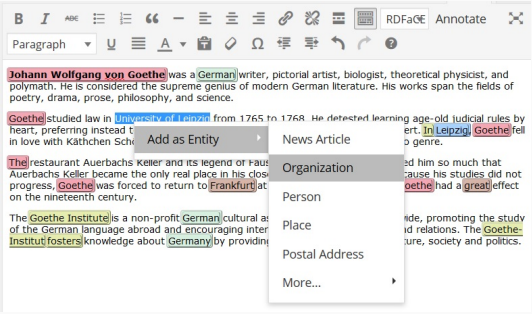

RDFaCE
Annotate
Paragraph
Johann Wolfgang von Goethe was a German writer, pictorial artist, biologist, theoretical physicist, and polymath. He is considered the supreme genius of modern German literature. His works span the fields of poetry, drama, prose, philosophy, and science.
Goethe studied law in University of Leipzig from 1765 to 1768. He detested learning age-old judicial rules by heart, preferring instead t ert. In Leipzig, Goethe fell in love with Käthchen Sch o genre.
Add as Entity
News Article
Organization
Person
Place
Postal Address
More...
The restaurant Auerbachs Keller and its legend of Faus d him so much that Auerbachs Keller became the only real place in his clos cause his studies did not progress, Goethe was forced to return to Frankfurt at oethe had a great effect on the nineteenth century.
The Goethe Institute is a non-profit German cultural as ide, promoting the study of the German language abroad and encouraging inter d relations. The Goethe-Institut fosters knowledge about Germany by providin ure, society and politics.

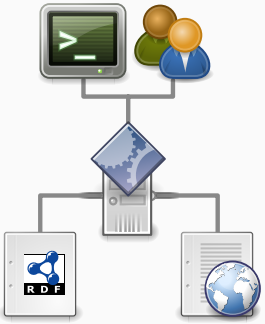
Problem

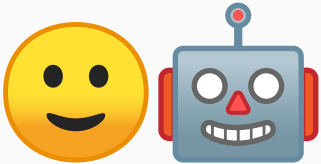

RDFaCE
Annotate
Paragraph
Johann Wolfgang von Goethe was a German writer, pictorial artist, biologist, theoretical physicist, and polymath. He is considered the supreme genius of modern German literature. His works span the fields of poetry, drama, prose, philosophy, and science.
Goethe studied law in University of Leipzig from 1765 to 1768. He detested learning age-old judicial rules by heart, preferring instead t
ert. In Leipzig, Goethe fell in love with Käthchen Sch
o genre.
The restaurant Auerbachs Keller and its legend of Faus
d him so much that Auerbachs Keller became the only real place in his clos
cause his studies did not progress, Goethe was forced to return to Frankfurt at
oethe had a great effect on the nineteenth century.
The Goethe Institute is a non-profit German cultural as
ide, promoting the study of the German language abroad and encouraging inter
d relations. The Goethe-Institut fosters knowledge about Germany by providin
ure, society and politics.
Add as Entity
News Article
Organization
Person
Place
Postal Address
More...



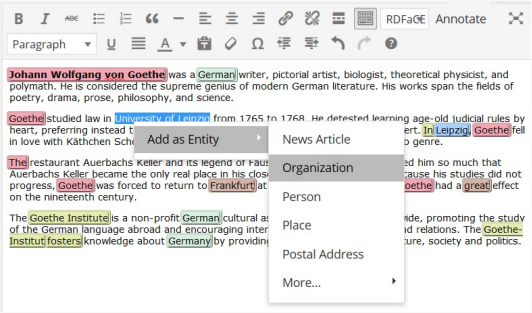
Problem

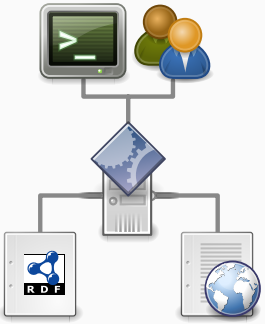

RDFaCE
Annotate
Paragraph
Johann Wolfgang von Goethe was a German writer, pictorial artist, biologist, theoretical physicist, and polymath. He is considered the supreme genius of modern German literature. His works span the fields of poetry, drama, prose, philosophy, and science.
Goethe studied law in University of Leipzig from 1765 to 1768. He detested learning age-old judicial rules by heart, preferring instead t
ert. In Leipzig, Goethe fell in love with Käthchen Sch
o genre.
Add as Entity
News Article
Organization
Person
Place
Postal Address
More...
The restaurant Auerbachs Keller and its legend of Faus
d him so much that Auerbachs Keller became the only real place in his clos
cause his studies did not progress, Goethe was forced to return to Frankfurt at
oethe had a great effect on the nineteenth century.
The Goethe Institute is a non-profit German cultural as
ide, promoting the study of the German language abroad and encouraging inter
d relations. The Goethe-Institut fosters knowledge about Germany by providin
ure, society and politics.

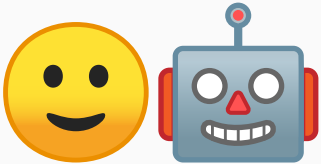

RDF

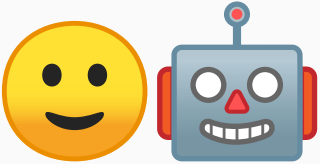


Problem

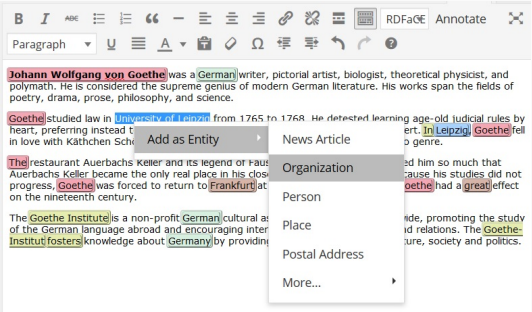

RDFaCE
Annotate
Paragraph
Johann Wolfgang von Goethe was a German writer, pictorial artist, biologist, theoretical physicist, and polymath. He is considered the supreme genius of modern German literature. His works span the fields of poetry, drama, prose, philosophy, and science.
Goethe studied law in University of Leipzig from 1765 to 1768. He detested learning age-old judicial rules by heart, preferring instead t
ert. In Leipzig, Goethe fell in love with Käthchen Sch
o genre.
The restaurant Auerbachs Keller and its legend of Faus
d him so much that Auerbachs Keller became the only real place in his clos
cause his studies did not progress, Goethe was forced to return to Frankfurt at
oethe had a great effect on the nineteenth century.
The Goethe Institute is a non-profit German cultural as
ide, promoting the study of the German language abroad and encouraging inter
d relations. The Goethe-Institut fosters knowledge about Germany by providin
ure, society and politics.
Add as Entity
News Article
Organization
Person
Place
Postal Address
More...

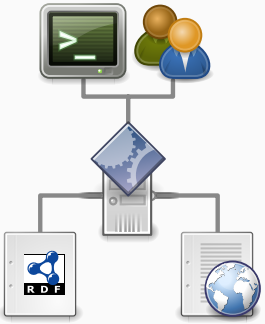

R D F

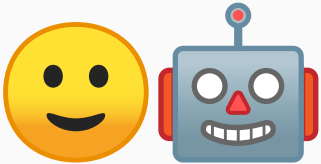




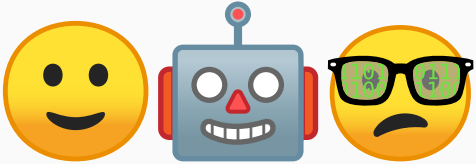
Problem



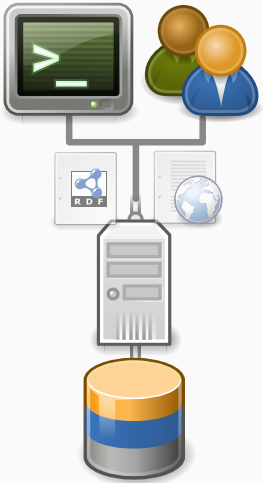


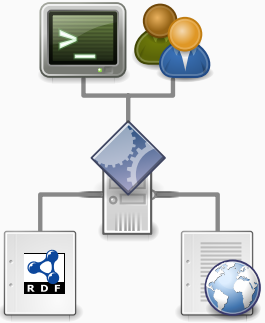


- hard to maintain

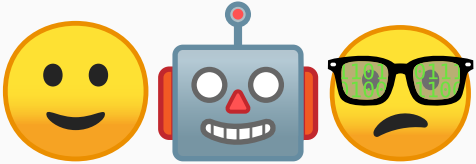


Problem

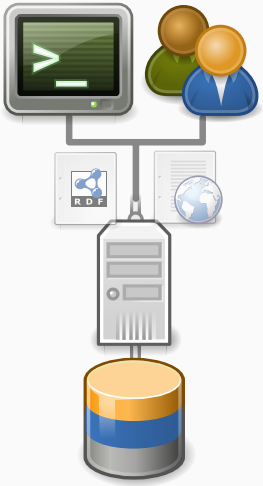


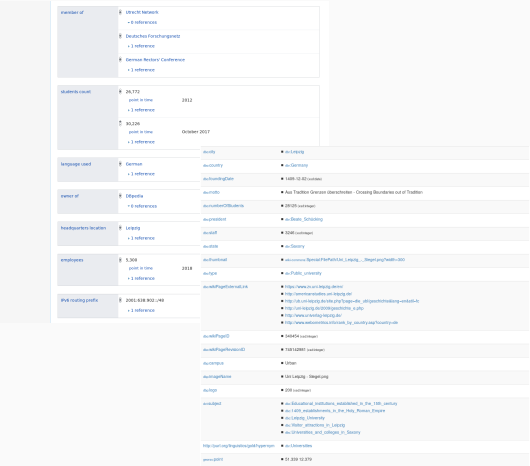


- hard to maintain



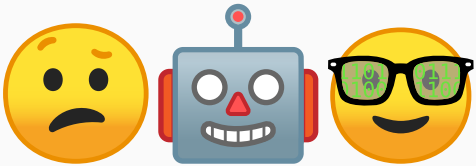
Problem

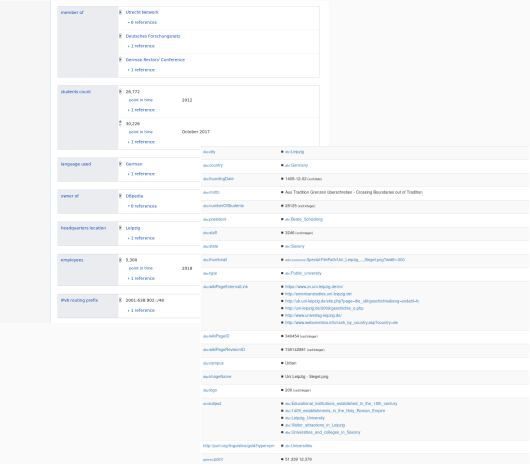

R D F


member of
Utrecht Network
• 6 references
Deutsches Forschungsnetz
• 1 reference
German Rectors' Conference
• 1 reference
students count
26,772
point in time
1912
• 1 reference
30,226
point in time
October 2017
• 1 reference
language used
German
• 1 reference
owner of
DBpedia
• 6 references
headquarters location
Leipzig
• 1 reference
employees
5,308
point in time
1918
• 1 reference
IPv6 routing prefix
2001:638:902::/48
• 1 reference
dbo:city
dbr:Leipzig
dbo:country
dbr:Germany
dbo:foundingDate
1409-12-02 (xsd:date)
dbo:motto
Aus Tradition Grenzen überschreiten - Crossing Boundaries out of Tradition
dbo:numberOfStudents
28125 (xsd:integer)
dbo:president
dbr:Beate_Schücking
dbo:staff
5246 (xsd:integer)
dbo:state
dbr:Saxony
dbo:thumbnail
en.commons.fkpmedia/FilePath/Uni_Leipzig_-_Siegel.png?width=300
dbo:type
dbr:Public_university
dbo:wikiPageExternalLink
https://www.bv.uni-leipzig.de/en/
http://amerikanistenliste.uni-leipzig.de/
http://ub.uni-leipzig.de/ebk.php?page=ifw_stil/geschichtsatlas-einstieb-fc
http://uni-leipzig.de/2009/geschichte_e.php
http://www.wi-wirtschaft-leipzig.de/
http://www.weltcomlinks.info/suche_by_country.asp?country=de
dbo:wikiPageID
345454 (xsd:integer)
dbo:wikiPageRevisionID
726132897 (xsd:integer)
dbo:campus
dbr:Urban
dbo:imageName
Uni Leipzig - Siegel.png
dbo:logo
208 (xsd:integer)
owl:sameAs
dbr:Educational_institutions_established_in_the_14th_century
dbr:1409_establishments_in_the_Holy_Roman_Empire
dbr:Leipzig_University
dbr:Walter_attractions_in_Leipzig
dbr:Universities_and_colleges_in_Saxony
http://purl.org/linguistics/gold/hypernym
dbr:Universities
geo:lat
51.339 12.379

Problem

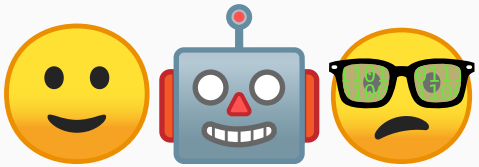
- Dynamic presentations are often not tailored to the content

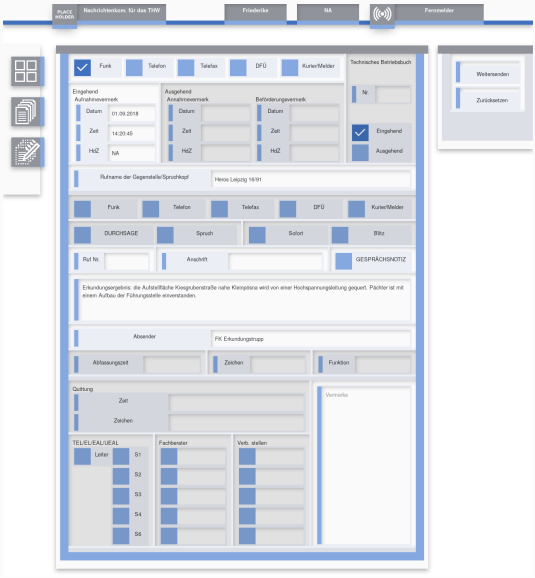




Problem

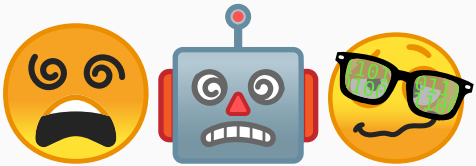
- Dynamic presentations are often not tailored to the content
- ... or they are built for a specific purpose

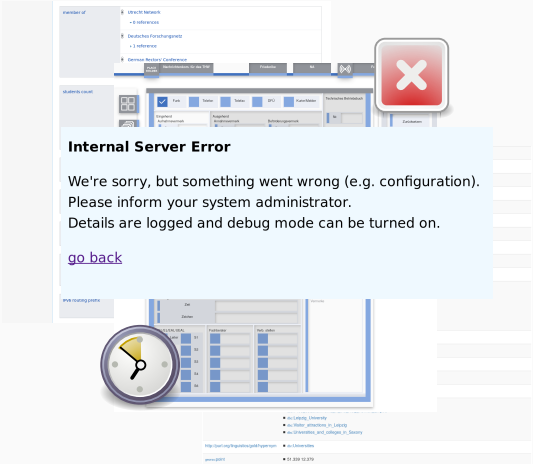




Problem

- Dynamic presentations are often not tailored to the content
- ... or they are built for a specific purpose
- We know SPARQL endpoints are not always reliable and break





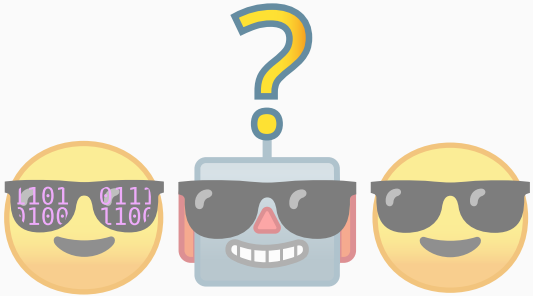
Problem

> *The Web was designed as an information space, with the goal not only that it should be useful for human–human communication, but also that machines would be able to participate and help users communicate with each other.*

Berners-Lee, T., Hendler, J.: Publishing on the semantic web. Nature 410(6832) (2001)

Problem

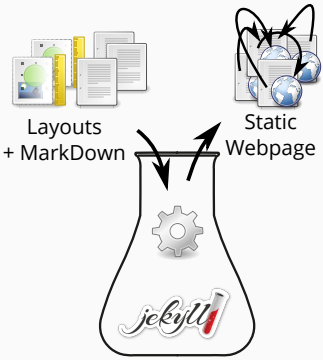
How can we close the gap between
structured data and human accessible exploration interfaces?



Approach

Approach

- We take the popular Jekyll system[1]
- It is a *Static Site Generator* and CMS[2]
- From a set of layout templates and markdown files it builds a webpage
- The webpage is static and self contained and thus can be served with minimum effort using e.g. nginx, github pages, static webspace
- Powerful, extensible, and simple templating language



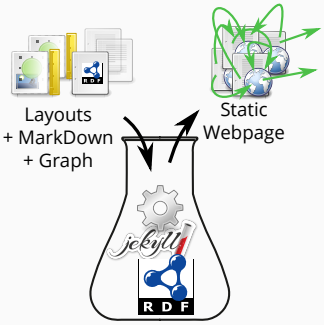
[1] https://jekyllrb.com/, https://www.staticgen.com/, https://www.netlify.com/blog/2016/05/02/top-ten-static-website-generators/

[2] Content Management System

Approach

A Jekyll plugin to include RDF data in your static site or build a complete site for your RDF graph

- Creates one page per resource
- Assign layouts per RDF class
- Standard simple Jekyll layouts
- + filters to add resource properties
- + filter to query the entire graph



Approach

```
<> a foaf:Document, schema:Event ;
  rdfs:label "7. Leipziger Semantic Web Tag"@de, "7th Leipzig Semantic Web Day"@en ;
  rdfs:comment """Der Leipziger Semantic Web Tag (LSWT) … [Language, Data and Knowledge (LDK) Konferenz](http://2019.ldk-conf.org/) …"""@de,
    """Since 2009, the Leipzig Semantic Web Day (LSWT) … [Language, Data and Knowledge (LDK) conference](http://2019.ldk-conf.org/) …"""@en;
  schema:logo </images/logo.png> ;
  schema:startDate "2019-05-22"^^xsd:date ;
  schema:endDate "2019-05-22"^^xsd:date ;
  schema:location lswt:Location ;
  site:registrationPage <https://pretix.eu/infai/lswt2019/> ;
  schema:organizer lswt:Team ;
  schema:partner lswt:Partner ;
  schema:subEvent lswt:StudentProgram ;
  schema:successorOf lswt:PreviousEvents .

lswt:Location a schema:PostalAddress ;
  rdfs:label "Veranstaltungsort"@de, "Venue"@en ;
  rdfs:comment "Das INNSIDE … gegenüber der Thomaskirche."@de, "The INNSIDE … just opposite to famous Thomaskirche."@en ;
  rdfs:seeAlso <https://www.melia.com/de/hotels/deutschland/leipzig/innside-leipzig/index.htm> ;
  schema:name "INNSIDE Hotel Leipzig" ;
  schema:streetAddress "Gottschedstraße 1" ;
  schema:postalCode "04109" ;
  schema:addressLocality "Leipzig" ;
```

RDF + HTML Template

```
<h1 class="display-4">{{ rootResource | rdf_property: "rdfs:label", page.language }}</h1>
{{ rootResource | rdf_property: "rdfs:comment", page.language | markdownify }}

<p><a href="{{ rootResource | rdf_property: "site:registrationPage" }}" role="button" class="btn btn-success">
  {{ "site:registrationPage" | rdf_property: "rdfs:label", page.language }}
</a></p>

<dl>
  <dt>{{ "schema:startDate" | rdf_property: "rdfs:label", page.language }}</dt>
  <dd>{{ rootResource | rdf_property: "schema:startDate" | localize: "%-d. %B %Y", page.language }}</dd>
  <dt>{{ "schema:location" | rdf_property: "rdfs:label", page.language }}</dt>
  {% assign location = rootResource | rdf_property: "schema:location" %}
  <dd><a href="{{ location.page_url }}">{% include address.html location=location %}</a></dd>
</dl>
```

Approach

```
<> a foaf:Document, schema:Event ;
  rdfs:label "7. Leipziger Semantic Web Tag"@de, "7th Leipzig Semantic Web Day"@en ;
  rdfs:comment """Der Leipziger Semantic Web Tag (LSWT) … [Language, Data and Knowledge (LDK) Konferenz](http://2019.ldk-conf.org/) …"""@de,
    """Since 2009, the Leipzig Semantic Web Day (LSWT) … [Language, Data and Knowledge (LDK) conference](http://2019.ldk-conf.org/) …"""@en;
  schema:logo </images/logo.png> ;
  schema:startDate "2019-05-22"^^xsd:date ;
  schema:endDate "2019-05-22"^^xsd:date ;
  schema:location lswt:Location ;
  site:registrationPage <https://pretix.eu/infai/lswt2019/> ;
  schema:organizer lswt:Team ;
  schema:partner lswt:Partner ;
  schema:subEvent lswt:StudentProgram ;
  schema:successorOf lswt:PreviousEvents .

lswt:Location a schema:PostalAddress ;
  rdfs:label "Veranstaltungsort"@de, "Venue"@en ;
  rdfs:comment "Das INNSIDE … gegenüber der Thomaskirche in Deutschland/Leipzig/Inside-Leipzig/index.htm…" ;
  rdfs:seeAlso <https://www.melia.com/de/hotels/deutschland/leipzig/inside-leipzig/index.htm> ;
  schema:name "INNSIDE Hotel Leipzig" ;
  schema:streetAddress "Gottschedstraße 1" ;
  schema:postalCode "04109" ;
  schema:addressLocality "Leipzig" ;
```

#LSWT2019

7th Leipzig Semantic Web Day

Since 2009, the Leipzig Semantic Web Day (LSWT) is an opportunity for scientists, enterprises, and organisations to discuss topics in the field of Semantic Technology. These emphases are demonstrated in presentations by experts from industry and science (schedule). If you have any further questions don't hesitate to contact the organisation team. The InfAI organizes the LSWT in conjunction with the Language, Data and Knowledge (LDK) conference (20-22 May) at the University of Leipzig or the Leipzig Semantic Web Day (LSWT) will be held this year at the media campus Villa Ida.

Registration

Date
22. May 2019

Place
INNSIDE Hotel Leipzig, Gottschedstraße 1, 04109 Leipzig

= Webpage

```
<h1 class="display-4">{{ rootResource | rdf_property: "rdfs:label", page.language }}</h1>
{{ rootResource | rdf_property: "rdfs:comment", page.language | markdownify }}

<p><a href="{{ rootResource | rdf_property: "site:registrationPage" }}" role="button" class="btn btn-success">
  {{ "site:registrationPage" | rdf_property: "rdfs:label", page.language }}
</a></p>

<dl>
  <dt>{{ "schema:startDate" | rdf_property: "rdfs:label", page.language }}</dt>
  <dd>{{ rootResource | rdf_property: "schema:startDate" | localize: "%-d. %B %Y", page.language }}</dd>
  <dt>{{ "schema:location" | rdf_property: "rdfs:label", page.language }}</dt>
  {% assign location = rootResource | rdf_property: "schema:location" %}
  <dd><a href="{{ location.page_url }}">{% include address.html location=location %}</a></dd>
</dl>
```

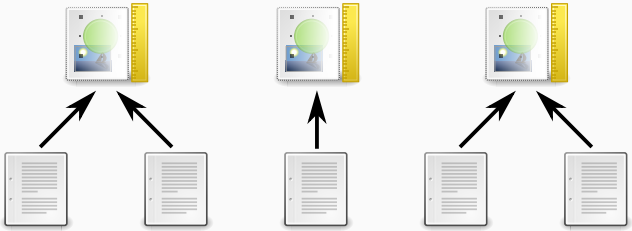
Approach

Template Mapping
Map Resources to Pages
Process Jekyll RDF Markup

Template Mapping

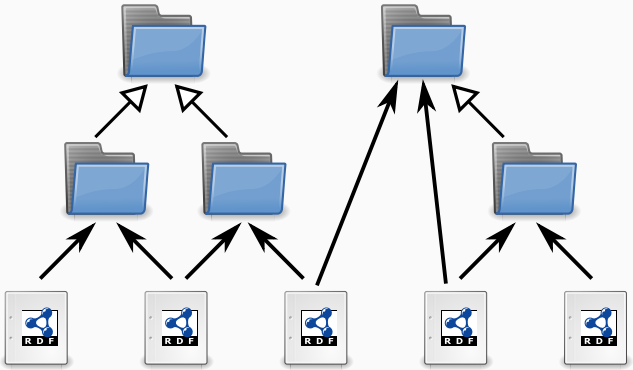
Template Mapping

Template Mapping

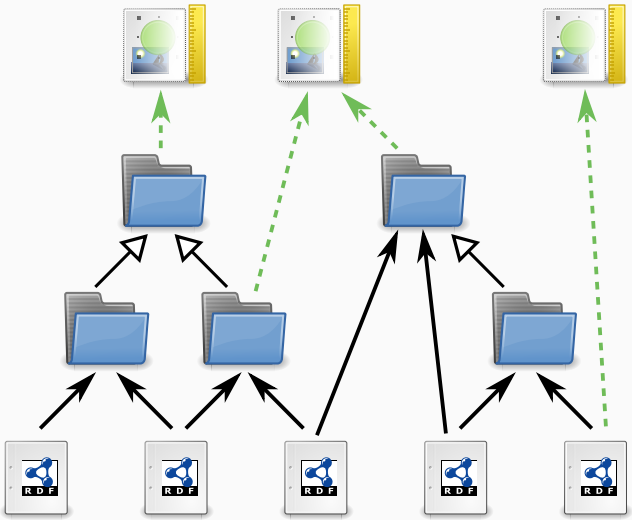


Template Mapping

Template Mapping



Template Mapping



Template Mapping

```
baseurl: "/sachsen/"
url: "http://pfarrerbuch.aksw.org"
plugins: [jekyll-rdf]
jekyll_rdf:
    path: "sachsen.ttl"
    restriction: "SELECT ?resourceUri WHERE {?resourceUri ?p ?o .
                  FILTER regex(str(?resourceUri),
                  '^http://pfarrerbuch.aksw.org/sachsen/')}"
    default_template: "resource"
    class_template_mappings:
        "http://xmlns.com/foaf/0.1/Person": "person"
        "http://purl.org/voc/hp/Place": "place"
        "http://purl.org/voc/hp/Position": "position"
    instance_template_mappings:
        "http://pfarrerbuch.aksw.org/": "home"
```

Resource Mapping and Site Creation

Resource Mapping and Site Creation

Resource Mapping and Site Creation

```
baseurl: "/sachsen/"
url: "http://pfarrerbuch.aksw.org"
plugins: [jekyll-rdf]
jekyll_rdf:
    path: "sachsen.ttl"
    restriction: "SELECT ?resourceUri WHERE {?resourceUri ?p ?o .
                  FILTER regex(str(?resourceUri),
                  '^http://pfarrerbuch.aksw.org/sachsen/')}"
    default_template: "resource"
    class_template_mappings:
        "http://xmlns.com/foaf/0.1/Person": "person"
        "http://purl.org/voc/hp/Place": "place"
        "http://purl.org/voc/hp/Position": "position"
    instance_template_mappings:
        "http://pfarrerbuch.aksw.org/": "home"
```

Resource Mapping and Site Creation

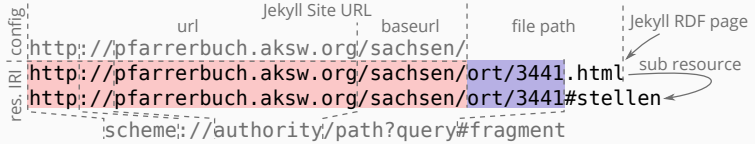


Figure 1: Scheme of the selection of page paths based on the configured Jekyll Site URL.

Jekyll RDF Markup Language

Jekyll RDF Markup Language

Jekyll RDF Markup Language

```
<h1>{{ page.title }}</h1>
<div>{{ page.date | date: "%Y-%m-%d" }}</div>

<div>{{ content }}</div>
```

Table 1: The filters defined by Jekyll RDF and the tasks performed by them.

Filter	Parameters ([optional])	Description
`rdf_get`		Get a variable representing a resource from the RDF graph.
`rdf_property`	IRI, [language, bool]	Get a value of a property of a resource. If the last parameter is set to true an array is returned (line 8).
`rdf_inverse_property`	IRI, [bool]	Get the value of an inverse property.
`rdf_collection` and `rdf_container`	[IRI]	Get RDF collections and RDF containers from the RDF graph.
`sparql_query`	sparql query	Execute a SPARQL Query on the RDF graph, the passed value is bound to `?resourceUri` or to `?resourceUri_n` if an array is provided.

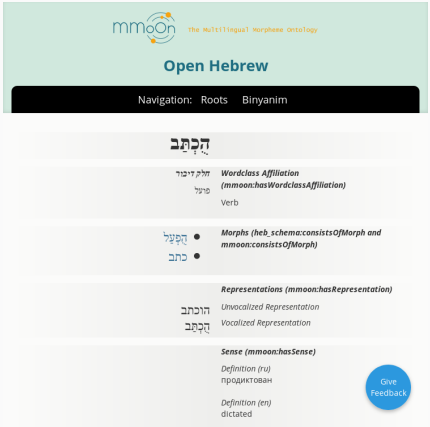
Jekyll RDF Markup Language

```
<h1>{{ page.rdf | rdf_property: "rdfs:label", "en" }}</h1>
<div>{{ page.rdf | rdf_property: "dct:created" | date: "%Y-%m-%d" }}</div>

{% assign publicationlist = "ex:publicationlist" | rdf_container %}
<ul>
{% for pub in publicationlist %}
  <li>{{ pub | rdf_property: "dc:title" }}</li>
  <li>{{ pub | rdf_property: "dct:creator", false, true | join: ", " }}</li>
{% endfor %}
</ul>
```

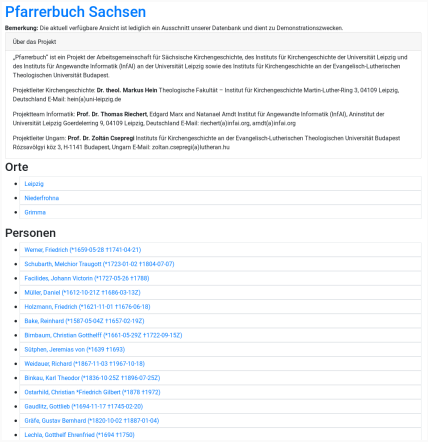
Examples

Examples



Setup	Open Hebrew Inventory
#Triples	197,374
#Templates/#Pages	4/2
#Res. Pages	13,426
Avg. Runtime	1952.4 s (32.54 min)

Examples



Pfarrerbuch Sachsen

Bemerkung: Die aktuell verfügbare Ansicht ist lediglich ein Ausschnitt unserer Datenbank und dient zu Demonstrationszwecken.

Über das Projekt

„Pfarrerbuch" ist ein Projekt der Arbeitsgemeinschaft für Sächsische Kirchengeschichte, des Instituts für Kirchengeschichte der Universität Leipzig und des Instituts für Angewandte Informatik (InfAI) an der Universität Leipzig sowie des Instituts für Kirchengeschichte an der Evangelisch-Lutherischen Theologischen Universität Budapest.

Projektleiter Kirchengeschichte: **Dr. theol. Markus Hein** Theologische Fakultät – Institut für Kirchengeschichte Martin-Luther-Ring 3, 04109 Leipzig, Deutschland E-Mail: hein(a)uni-leipzig.de

Projektteam Informatik: **Prof. Dr. Thomas Riechert**, Edgard Marx and Natanael Arndt Institut für Angewandte Informatik (InfAI), Anninstitut der Universität Leipzig Goerdelering 9, 04109 Leipzig, Deutschland E-Mail: riechert(a)infai.org, arndt(a)infai.org

Projektleiter Ungarn: **Prof. Dr. Zoltán Csepregi** Institutes für Kirchengeschichte an der Evangelisch-Lutherischen Theologischen Universität Budapest Rózsavölgyi köz 3, H-1141 Budapest, Ungarn E-Mail: zoltan.csepregi(a)lutheran.hu

Orte

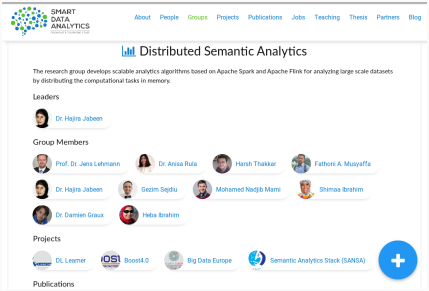
- Leipzig
- Niederfrohna
- Grimma

Personen

- Werner, Friedrich (*1659-05-28 †1741-04-21)
- Schubarth, Melchior Traugott (*1723-01-02 †1804-07-07)
- Facilides, Johann Victorin (*1727-05-26 †1788)
- Müller, Daniel (*1612-10-21Z †1686-03-13Z)
- Holzmann, Friedrich (*1621-11-01 †1676-06-18)
- Bake, Reinhard (*1587-05-04Z †1657-02-19Z)
- Birnbaum, Christian Gotthelff (*1661-05-29Z †1722-09-15Z)
- Stöphen, Jeremias von (*1639 †1693)
- Waldauer, Richard (*1867-11-03 †1967-10-18)
- Binkau, Karl Theodor (*1836-10-25Z †1896-07-25Z)
- Osterfeld, Christian *Friedrich Gilbert (*1878 †1972)
- Gaudlitz, Gottlieb (*1694-11-17 †1745-02-20)
- Gräfe, Gustav Bernhard (*1820-10-02 †1887-01-04)
- Lechla, Gotthelf Ehrenfried (*1694 †1750)

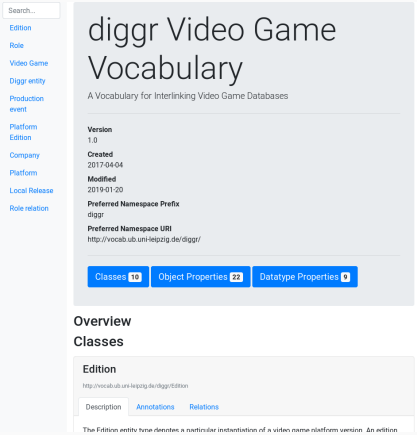
Setup	Pfarrerbuch (demo subset)
#Triples	1,685
#Templates/#Pages	4/-
#Res. Pages	138
Avg. Runtime	8 s

Examples



Setup	SDA Work Group
#Triples	27,295
#Templates/#Pages	8 (Class) + 49 (Instance)/2
#Res. Pages	253
Avg. Runtime	327.3 s (5.45 min)

Examples



Setup	diggr Vocabulary
#Triples	221
#Templates/#Pages	1/-
#Res. Pages	1
Avg. Runtime	8 s

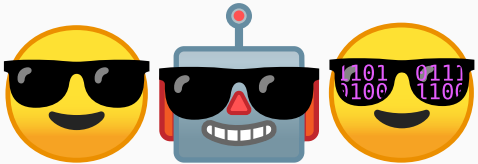
Conclusion

Conclusion

- Create webpages from RDF data sources (CMS, exploration interface)
- Cost efficient serving of static sites
- Customizability of the layout without code, by using a templating language

Conclusion

- Create webpages from RDF data sources (CMS, exploration interface)
- Cost efficient serving of static sites
- Customizability of the layout without code, by using a templating language



Hands On

Hands On

https://github.com/white-gecko/JekyllRDF-Tutorial

https://natanael.arndt.xyz/2018/08/07/jekyll-rdf-tutorial-screencast

https://vimeo.com/283525074

Contact

Contact

Natanael Arndt
arndt@informatik.uni-leipzig.de
https://natanael.arndt.xyz
twitter: @white_gecko
mastodon: @white_gecko@chaos.social

Web Page: http://aksw.org/Projects/JekyllRDF
GitHub: https://github.com/AKSW/jekyll-rdf

Jekyll RDF: Template-Based Linked Data Publication with Minimized Effort and Maximum Scalability Natanael Arndt, Sebastian Zänker, Gezim Sejdiu, Sebastian Tramp *in* Proceedings of the 19th International Conference on Web Engineering (ICWE 2019)
https://natanael.arndt.xyz/bib/arndt-n-2019--jekyll-rdf
https://doi.org/10.1007/978-3-030-19274-7_24

Credits

Credits

- Emojis licensed under Apache License Version 2.0
- 🙂 😐 😒 😕 🧐 😵 😎 🤓 🤖 Copyright 2013 Google, Inc. All Rights Reserved.
- 🤓 😎 🤓 😎 🤓 🤖 🤖 🤖 🕶️ 🤖 🤖 🤖 🚤 Copyright 2013 Google, Inc. All Rights Reserved, Copyright 2019 Natanael Arndt <arndtn@gmail.com> All Rights Reserved.

Credits

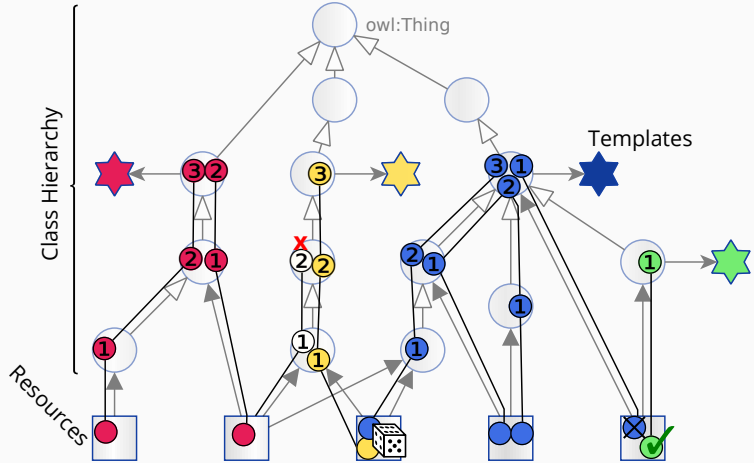


Figure 2: The class hierarchy is used to select the template to render a resource.